Control of Robotic Manipulation

The invention relates to control of robotic manipulation; in particular motion compensation in robotic manipulation. The invention further relates to the use of

5   of stereo images.

Robotic manipulation is known in a range of fields. Typical systems include a robotic manipulator such as a robotic arm which is remote controlled by a user. For example the robotic arm may be configured to mirror the actions of the

10   human hand. In that case a human controller may have sensors monitoring actions of the controller's hand. Those sensors provide signals allowing the robotic arm to be controlled in the same manner. Robotic manipulation is useful in a range of applications, for example in confined or in miniaturised/microscopic applications.

15

One known application of robotic manipulation is in medical procedures such as surgery. In robotic surgery a robotic arm carries a medical instrument. A camera is mounted on or close to the arm and the arm is controlled remotely by a medical practitioner who can view the operation via the camera. As a result

20   keyhole surgery and microsurgery can be achieved with great precision. A problem found particularly in medical procedures but also in other applications arises when it is required to operate on a moving object or moving surface such as a beating heart. One known solution in medical procedures is to hold the relevant surface stationary. In the case of heart surgery it is known to stop the

25   heart altogether and rely on other life support means while the operation is taking place. Alternatively the surface can be stabilised by using additional members to hold it stationary. Both techniques are complex, difficult and increase the stress on the patient.

One proposed solution is set out in US5971976 in which a position controller is also included. The medical instrument is mounted on a robotic arm and remotely controlled by a surgeon. The surface of the heart to be operated on is mechanically stabilised and the stabiliser also includes inertia or other

5   position/movement sensors to detect any residual movement of the surface. A motion controller controls the robotic arm or instrument to track the residual movement of the surface such that the distance between them remains constant and the surgeon effectively operates on a stationary surface. A problem with this system is that the arm and instrument are motion locked to a specific point

10   or zone on the heart defined by the mechanical stabiliser but there is no way of locking it to other areas. As a result if the surgeon needs to operate on another region of the surface then the residual motion will no longer be compensated and can indeed be enhanced if the arm is tracking another region of the surface, bearing in mind the complex surface movement of the heart.

15

The invention is set out in the appended claims. Because the motion sensor can sense motion of a range of points, the controller can determined the part of the object to be tracked. Eye tracking relative to a stereo image allows the depth of a fixation point to be determined.

20

Embodiments of the invention will now be described, by way of example, with reference to the drawings of which:

Fig. 1 is a schematic view of a known robotic manipulator;

25   Fig. 2 shows the components of an eye tracking system;

Fig. 3 shows a robotic manipulator according to the invention;

Fig. 4 shows a schematic view of a stereo image display; and

Fig. 5 shows the use of stereo image in depth determination.

Referring to Fig. 1 a typical arrangement for performing robotic surgery is shown designated generally 10. A robotic manipulator 20 includes an articulated arm 22 carrying a medical instrument 24 as well as the cameras 26. The arm is mounted on a controller 28. A surgical station designated generally

5   40 includes binocular vision eye pieces 42 through which the surgeon can view a stereo image generated by cameras 26 and control gauntlets 44. The surgeon inserts his hands into the control gauntlets and controls a remote analogue of the robotic manipulator 20 based on the visual feedback from eyepiece 42. Interface between the robotic manipulator 20 and surgical station 40 is via an

10   appropriate computer processor 50 which can be of any appropriate type for example a PC or laptop. The processor 50 conveys the images from camera 26 to the surgical station 40 and returns control signals from the robotic arm analogue controlled by the surgeon via gauntlets 44. As a result a fully fed back surgical system is provided. Such a system is available under the

15   trademark Da Vinci Surgical Systems from Intuitive Surgical, Inc of Sunnyvale California USA or Zeus Robotic Surgical Systems from Computer Motion, Inc Goleta California USA. In use the surgical instrument operates on the patient and the only incision required is sufficient to allow camera vision and movement of the instrument itself as a result of which minimal stress to the

20   patient is introduced. Furthermore using appropriate magnifications/reduction techniques, micro surgery can very easily take place.

As discussed above, it is known to add motion compensation to a system such as this whereby motion sensors on the surface send a movement signal which is

25   tracked by the robotic arm such that the surface and arm are stationary relative to one another. In overview the present invention further incorporates an eye tracking capability at the surgical station 40 identifying which part of the surface the surgeon is fixating on and ensuring that the robotic arm tracks that particular point, the motion of which may vary relative to other points because

4

of the complex motion of the heart's surface. As a result the invention achieves dynamic reference frame locking.

Referring to Fig. 2 an appropriate eye tracking arrangement is shown
5   schematically. The user 60 views an image 62 on a display 63. An eye-tracking device 70 includes one or more light projectors 71 and a light detector 72. In practice the light projectors may be infra-red (IR) LEDs and the detector may be an IR camera. The LEDs project light 73 onto the eye of the user 60 and the angle of gaze of the eye can be derived using known techniques by
10   detecting the light 74 reflected onto the camera. Any appropriate eye tracking system may in practice be used for example an ASL model 504 remote eye-tracking system (Applied Science Laboratories, MA, USA). This embodiment may be particularly applicable when a single camera is provided on the articulated arm 22 of a robotic manipulator and thus a single image is presented
15   to the user. The gaze of the user is used to determine the fixation point of the user on the image 62. It will be appreciated that a calibration stage may be incorporated on initialisation of any eye-tracking system to accommodate differences between users' eyes or vision. The nature of any such calibration stage will be well known to the skilled reader.
20

Referring now to Fig. 3, the robotic arm and tracking system are shown in more detail.

An object 80 is operated on by a robotic manipulator designated generally 82.
25   The manipulator 82 includes 3 robotic arms 84, 86, 88 articulated in any appropriate manner and carrying appropriate operating instruments. Arm 84 and arm 86 each support a camera 90a, 90b displaced from one another sufficient to provide stereo imaging according to known techniques. Since the

relative positions of the three arms are known, the position of the cameras in 3D space is also known.

5 In use the system allows motion compensation to be directed to the point on which the surgeon is fixating (i.e. the point he is looking at, at a given moment). Identifying the fixation point can be achieved using known techniques which will generally be built in with an appropriate eye tracking device provided, for example, in the product discussed above. In the preferred embodiment the cameras are used to detect the motion of the fixation point and

10 send the information back to the processor for control of the motion of the robotic arm.

In particular, once, at any one moment, the fixation point position is identified on the image viewed by the human operator, given that the position of the

15 stereo cameras 90a and 90b are known the position of the point on the object 80 can be identified. Alternatively, by determining the respective direction of gaze of each eye, this can be replicated at the stereo camera to focus on the relevant point. The motion of that point is then determined by stereo vision. In particular, referring to Fig. 5 it will be seen that the position of a point can be

20 determined by measuring the disparity in the view taken by each camera 90a, 90b. For example for a relatively distant object 100 on a plane 102 the cameras take respective images A1, B1 defining a distance X1. A more distant object 104 creates images A2, B2 in which the distance between the objects as shown in the respective images is X2. There is an inverse relationship between the

25 distance and the depth of the point. As a result the relative position of the point to the camera can be determined.

In particular, the computer 50 calculates the position in the image plane of the co-ordinates in the real world (so-called "world coordinates"). This may be done as follows:

5   A 3D point $M = [x,y,z]^T$ is projected to a 2D image point $m = [x,y,]^T$ through a 3x4 projection matrix P, such that $S\dot{m} = P\dot{M}$, where S is a non-zero scale factor and $\dot{m} = [x, y, 1]^t$ and $\dot{M} = [x,y,z,1]^t$. In binocular stereo systems, each physical point M in 3D space is projected to m_1 and m_2 in the two image planes, i.e;

$$S_1 \dot{m}_1 = P_1 \dot{M}$$

10     $$S_2 \dot{m}_2 = P_2 \dot{M} \qquad\qquad\qquad (1)$$

If we assume that the world coordinate system is associated with the first camera, we have

15     $$P_1 = [A|0]$$
       $$P_2 = [A'R|A' t] \qquad\qquad\qquad (2)$$

Where R and t represent the 3x3 rotation matrix and the 3x1 translation vector defining the rigid displacement between the two cameras.

20

The matrices A and A^1 are the 3x3 intrinsic parameter matrices of the two cameras. In general, when the two cameras have the same parameter settings and with square pixels (aspect ration = 1), and the angle (θ) between the two image coordinate axes being $\pi/2$ we have:

25

$$A = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \qquad\qquad\qquad (3)$$

Where (u_0, v_0) are the coordinates of the image principal point, i.e, the point where points located at infinity in world coordinates are projected.

Generally, matrix A can have the form of

$$A = \begin{bmatrix} f_u & f_u \cot\theta & u_0 \\ 0 & f_v / \sin\theta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

Where f_u and f_v correspond to the focal distance in pixels along the axes of the image. All parameters of A can be computed through classical calibration method (e.g. as described in the book by O. Faugeras, "Three-Dimensional Computer Vision: a Geometric Viewpoint", MIT press, Cambridge, MA, 1993).

Known techniques for determining the depth are for example as follows. Firstly, the apparatus is calibrated for a given user. The user looks at predetermined points on a displayed image and the eye tracking device tracks the eye(s) of the user as they look at each predetermined point. This sets the user's gaze within a reference frame (generally two-dimensional if one image is displayed and three-dimensional if stereo images are displayed). In use, the user's gaze on the image(s) is tracked and thus the gaze of the user within this reference frame is determined. The robotic arms 84, 86 then move the cameras 90a, 90b to focus on the determined fixation point.

For instance, consider Figure 2 again which shows a user 60, an image 62 on a display 63 and an eye tracking device 70. In use, the tracking device 70 is first calibrated for the user. This involves the computer 50 displaying on the display

8

a number of pre-determined calibration points, indicated by 92. A user is instructed to focus on each of these in turn (for instance, the computer 50 may cause each calibration point to be displayed in turn). As the user stares at a calibration point, the eye tracking device 70 tracks the gaze of the user. The

5   computer then correlates the position of the calibration point with the position of the user's eye. Once all the calibration points have been displayed to a user and the corresponding eye position recorded, the system has been calibrated to the user.

10   Subsequently a user's gaze can be correlated to the part of the image being looked at by the user. For each eye, the coordinates $[x_l, y_l]$ and $[x_r, y_r]$ are known from each eye tracker from which $[x, y, z]^T$ can be calculated from Equations (1)-(4).

15   By carrying out this step across time the motion of the point fixated on by the human operator can be tracked and the camera and arm moved by any appropriate means to maintain a constant distance from the fixation point. This can either be done by monitoring the absolute position of the two points and keeping it constant or by some form of feedback control such as using PID

20   control. Once again the relevant techniques will be well known to the skilled person.

It will be further recognised that the cameras can be focussed or directed towards the fixation point determined by eye-tracking, simply by providing

25   appropriate direction means on or in relation to the robotic arm. As a result the tracked point can be moved to centre screen if desired.

In the preferred embodiment the surgical station provides a stereo image via binocular eyepiece 42 to the surgeon, where the required offset left and right images are provided by the respective cameras mounted on the robotic arm.

5   According to a further aspect of the invention enhanced eye tracking in relation to stereo images is provided. Referring to Fig. 4, a further embodiment of the invention is shown. The system requires left and right images slightly offset to provide, when appropriately combined, a stereo image as well known to the skilled reader. Images of a subject being viewed are displayed on displays

10   200a, 200b. These displays are typically LCD displays. A user views the images on the displays 200a, 200b through individual eye pieces 202a, 202b via intermediate optics including mirrors 204a, b, c (and any appropriate lens although any appropriate optics can of course be used).

15   Eye tracking devices are provided for each individual eye piece. The eye-tracking device includes light projectors 206 and light detectors 208[a,b]. In a preferred implementation , the light projectors are IR LEDs and the light detector comprises an IR camera for each eye. An IR filter may be provided in front of the IR camera. The images (indicated in Figure 4 by the numerals

20   210a, 210b) captured by the light detectors 208a, 208b show the position of the pupils of each eye of the user and also the Purkinje Reflections of the light sources 206.

The angle of gaze of the eye can be derived using known techniques by the

25   detecting the reflected light.

In a preferred, known implementation Purkinje images are formed by light reflected from surfaces in the eye. The first reflection takes place at the anterior surface of the cornea while the fourth occurs at the posterior surface of the lens

of the eye. Both the first and fourth Purkinje images lie in approximately the same plane in the pupil of the eye and, since eye rotation alters the angle of the IR beam from the IR projectors 206 with respect to the optical axis of the eye, and eye translations move both images by the same amount, eye movement can be obtained from the spatial position and distance between the two Purkinje reflections. This technique is commonly known as the Dual-Purkinje Image (DPI) technique. DPI also allows for the calculation of a user's accommodation of focus i.e. how far away the user is looking. Another eye tracking technique subtracts the Purkinje reflections from the nasal side of the pupil and the temporal side of the pupil and uses the difference to determine the eye position signal. Any appropriate eye tracking system may in practice be used for example an ASL model 504 remote eye-tracking system (Applied Science Laboratories, MA, USA).

By tracking the individual motion of each eye and identifying the fixation point F on the left and right images 200a, 200b, not only the position of the fixation point in the X Y plane (the plane of the images) can be identified but also the depth into the image, in the Z direction.

Once the eye position signal is determined, the computer 50 uses this signal to determine where, in the reference field, the user is looking and calculates the corresponding position on the subject being viewed. Once this position is determined, the computer signals the robotic manipulator 82 to move the arms 84 and/or 86 which support the cameras 90a and 90b to focus on the part of the subject determined from the eye-tracking device, allowing the motion sensor to track movement of that part and hence lock the frame of reference to it.

Although the invention has been described with reference to eye tracking devices that use reflected light, other forms of eye tracking may be used, e.g.

11

measuring the electric potential of the skin around the eye(s) or applying a special contact lens and tracking its position.

It will be appreciated that the embodiments above and elements thereof can be combined or interchanged as appropriate. Although specific discussion is made of the application of the invention to surgery, it will be recognised that the invention can be equally applied in many other areas where robotic manipulation or stereo imaging is required. Although stereo vision is described, monocular vision can also be applied. Also other appropriate means of motion sensing can be adopted, for instance, by the use of casting structured light onto the object and observing changes as the object moves, or by using laser range finding. These examples are not supposed to be limiting.